

Technologies for the study of speech: Review and an application

Elena Babatsouli
ebabatsouli@ismbs.eu

Institute of Monolingual and Bilingual Speech, Crete, Greece

Abstract. Technologies used for the study of speech are classified here into non-intrusive and intrusive. The paper informs on current non-intrusive technologies that are used for linguistic investigations of the speech signal, both phonological and phonetic. Providing a point of reference, the review covers existing technological advances in language research methodology for normal and disordered speech, cutting across boundaries of sub-fields: child vs. adult; first, second, or bilingual; standard or dialectal. The indispensability of such technologies for investigating linguistic aspects of speech is evident in an application that uses CLAN to manage and perform phonological and computational analysis of a bilingual child's dense speech data, longitudinally over 17 months of phonological development in Greek/English. Backed by acoustic analysis using Praat, developmental results on the child's labio-velar approximant in English are also given.

Keywords: technologies, normal/disordered speech, child development, CLAN

Introduction

Speech data feed the observation, documentation and analysis of oral language; theoretical speculations in phonology are grounded on phonetic evidence which, in turn, provide yardsticks for comparing the production of speech cross-linguistically in all its likely forms: child and adult; standard or dialectal; first (L1), second (L2), bilingual (L1'2); normal or disordered. Despite the presence of universal patterns (Jakobson, 1941/1968) in speech, production variability is a prevalent phenomenon, sometimes going unnoticed and at other times drawing the listener's attention to inaccuracy and even unintelligibility, as witnessed in random errors, idiolects, dialects, accents, speech disorders, and speech impairments. Detailed investigations of such productions, typical or atypical, provide insights into the overt and hidden processes involved in the human linguistic faculty, namely, perception, processing, and articulation. Different methodologies target investigation into separate components of the faculty of speech, some focusing on production both from the listener's point of view (e.g. metalinguistic knowledge, acoustics etc.), or the speaker's point of view (e.g. articulation), and others focusing on perception, and aspects of cognitive processing.

Thus, the significance of efficient and systematic evaluation that encompasses reliable and uniform documentation and analysis of speech data in all their forms becomes paramount; this is tackled in the field of experimental phonetics with the use of instrumentation and technologies: the collection of instruments, techniques and processes that facilitate the measurement, monitoring and analysis of speech, both in terms of production variables and the processes leading to them; such instrumentation necessitates the use of machines, computers, software, and other devices. It is not unusual to find compartmentalized use of technologies in research methodologies used even across sub-fields (L1, L2, L1'2, SSDs) of the same discipline (phonology/phonetics).

A holistic approach is used in this paper that reviews prevalent approaches and technological means for the study of normal and disordered speech, following a brief historical note. Technologies for the study of speech are classified here into non-intrusive (NITs) and intrusive (InT), referring to whether the analyses target the speech signal alone (thus, non-intrusive to the speaker per se), or the analyses rely on the physical participation of the speaker in the experiment (thus, intrusive). NITs utilized in the study of speech generally can, and do have direct applications on speech sound disorders (SSDs), and vice versa. To our knowledge, there is a lack of reviews on NITs. On the other hand, extensive discussions of intrusive technologies have been addressed elsewhere (see section on InTs below), so only a brief mention is made here for inclusiveness. More than a review, the paper provides an example of non-intrusive technology in use: CLAN (MacWhinney, 2000) is utilized to manage a bilingual toddler's longitudinal data over 17 months of phonological development in Greek and English. The aim is to automate computations for the establishment of developmental paths in consonantal production. Specifically, results on the development of the child's labio-velar approximant in English are given, supported by acoustic analysis using Praat (Boersma & Weenink, 2015).

The paper is organized as follows: non-intrusive technologies (reproducing the speech signal, transcription, speech analysis software); technologies; NITs in use: a case study; conclusions.

Non-intrusive technologies

This section introduces methods and technologies for the study of language in general, and for phonology/phonetics in particular, by focusing on the speech signal (what reaches the listener's ear) rather than on targeting the motor, sensory, perceptual, and cognitive mechanisms behind the production of speech. Technology supports and validates the researcher's analysis and evaluation. The majority of technologies presented in this section are freely available online, and they come with manuals to assist self-instruction.

Reproducing the speech signal

It is not until the invention of the phonograph in the 19th century that experimental phonetics moved from "playing it by ear" to recording and reproducing the speech signal. This has greatly assisted phoneticians who are permitted to replay the speech signal as needed and document it in detailed annotation. Quality digital recording is nowadays easily available, even at the touch of a button on a smartphone. However, considerations of the role of environmental noise on the quality of the resulting sound for the purpose of research necessitate a sound-proof environment, where more specialized acoustic software supplemented by microphone(s) may be employed, such as the widely used Praat (Boersma & Weenink, 2015), and Sony's SoundForge. Digitized recordings in e.g. wav, aiff, mp3 formats are usually fully compatible with most standard computers, or they can be easily converted to specific formats, following the requirements of individual computer software, using online sound converters, such as the free software SoundConverter.

Phonetic transcription

First rudimentary attempts at manually annotating speech are found in the form of phonetic transcription of Sanskrit in 500 B.C. (Allen, 1953). In modern times, Alexander Melville Bell's Visible Speech (1867) is an early forerunner of the International Phonetic Association (IPA, 1999) chart (graphemes, diacritics, and suprasegmentals) that provides a notational system for speech sounds across languages. IPA was principally designed for the annotation of

normal speech, but it was subsequently extended with additional symbols (ExtIPA) to account for disordered speech (Duckworth et al., 1990). While IPA is more widely used worldwide, other systems of transcription (including use of the standard alphabet) also exist, such as the American Phonetic Alphabet (APA) which utilizes a combination of standard graphemes (from Latin or Greek) and diacritics but no uniquely designed symbols. An alternate notational system for clinical phonetics, SK, proposed by Shriberg & Kent (2003; 2013) is intended to inform on the targeted as well as the produced sound, rather than just the production. The reliability of this notational system and the possibility of combining SK and IPA are discussed in Ball (2008). It is widely admitted, nevertheless, that phonetic transcription alone, even in its narrowest form (Ball et al., 2009) is not an adequate measure of produced speech because of its “impressionistic” (predicated on impression) (Abercrombie, 1967; Hayward, 2014) and, thus, subjective nature, evident in different denotational values cross-linguistically (Edwards & Munson, 2012).

Automating phonetic transcription

IPA symbols in computer-readable format are found in Unicode (ISO 10646). Other ASCII phonetic scripts include Arpabet, TIMITBET, SAM-PA (Speech Assessment and Methods: Phonetic Alphabet) and X-SAMPA (Extended SAMPA). It is possible to type phonetic symbols directly into computer files using downloadable IPA keyboards created by Keyman (http://scripts.sil.org/cms/scripts/page.php?site_id=nrsi&id=uniipakeyboard). Computer programs, like TranscriberAG (Barras et al., 2001), are designed to assist the manual annotation of speech signals by segmenting long speech recordings, transcribing and labeling various aspects of speech (e.g. turns, topic changes, acoustic conditions). SIL's Phonology Assistant (http://www.sil.org/resources/software_fonts/phonology-assistant) automatically charts phones of a given phonetic corpus, permitting fine tuning of transcription and phonetic parsing. Complex annotations of video and audio files on many tiers can be made in several computer programs, like Max Planck Institute's EUDICO Linguistic Annotator (ELAN, 2015), which allows easy navigation and searches through both the media and the annotations (Sloetjes & Wittenburg, 2008). It is well-known, however, that narrow phonetic transcription, even when done on a computer, is formidable and very labor intensive (Schmid, Verspoor & MacWhinney, 2011).

Automatic Speech Recognition (ARS), which is the computerized rendition of spoken words into text, has been utilized towards easing the task of manual transcription. WebMAUS: Automatic Segmentation and Labelling of Speech Signals over the Web, which partly utilizes ARS technology, may be found available online (IPS, 2014). The MAUS program (Schiel, Draxler & Harrington, 2011) time-aligns pre-uploaded recorded speech and its matching text in regular orthography, and produces a phonemic encoding of the standard pronunciation (in SAM-PA script); subsequently, a machine-learned expert system uses this phonemic encoding and creates a graph of all possible pronunciation variants for the targeted language which is aligned back to the corresponding sound file.

There are other ARS programs that also permit automatic speech annotation, such as Nuance's Dragon Naturally Speaking and Microsoft's Speech Recognition. Although these programs have moved the annotation task way beyond handwriting and have partly automated the process, they are still wanting in some aspects. While it is possible to annotate clear and precise dictation from faultless audio recordings into text formats, it is not yet possible to achieve full accuracy in the phonetic transcription of ordinary conversational speech. This is because of limitations with regard to: a) differentiating between diverse speaking styles, both isolated and in conversation, b) annotating running speech in spite of different word boundaries (e.g. it's a nice cream vs. it's an ice-cream) and

homophones (like pale and pail), and c) dealing with superfluous noise in the recordings that inhibits accurate acoustic analysis. Despite these drawbacks, higher transcription accuracy may eventually be achieved by training ARS programs to specific speech types. There is evidence of dissemination of ARS methods in the investigation of speech variability, as in the case of identifying L2 errors (Strik et al., 2009; Weinberger & Kunath, 2011), but the approach is still young. On the other hand, ARS programs are systematically used as assistive technology providing corrective feedback to learners to improve their pronunciation (Cucchiaroni, Neri & Strik, 2009) and for people with cognitive, physical or sensory impairments to enhance learning, and improve communication skills (Wald & Bain, 2008).

Speech Analysis Software (SAS): managing and retrieving data

The downside of trouble-free recordings has been that researchers end up with large amounts of speech data whose sheer volume hinders the task of research. Speech banks targeting and amassing specific types of data are being created around the world, even as we speak; up-to-date reviews of corpora have been addressed by Durand, Gut & Kristoffersen (2014) and Ruhi et al. (2014). Among the most established Speech Banks is the TalkBank project (MacWhinney, 2007) that includes the CHILDES database (speech records for child monolingual data); the PhonBank project (a database for the study of phonological development); BilingBank (bilingual corpora); and the AphasiaBank (multimedia interactions for the study of communication in aphasia). The Speech Accent Archive (Weinberger, 2015) chronicles hundreds of cross-linguistic examples of L2 English speech, as well as native English dialects. The LeaP corpus (Gut, 2012) is concerned with the acquisition of prosody by non-native speakers of German and English. Max Planck Institute's IMDI-corpora (2015) is a collection of spoken and written cross-linguistic data. Also, a Disordered Speech Bank is currently under way for the collection of examples of acquired SSDs across languages and in multilingual contexts (Ball, 2015). Cole et al. (2011:442) presented online depositories of phonological resources, like the UCLA Phonetics Laboratory Archive and VoxForge.

There is an increasing number of freely available computer software, like ELAN mentioned earlier, that are designed for the purpose of facilitating the creation of such speech databases and corpora (MacWhinney & Snow, 1985; Durand et al., 2014; Rose & MacWhinney, 2014), thus easing the management, organization, and retrieval of large data volumes. In these programs, audio or video extracts are time-aligned to orthographic and/or phonetic transcription, and may be supplemented by coding schemes, thus, permitting efficient and fast retrieval of specific information sets on the data entered. The use of coding schemes alongside phonetic annotation is important for 'recognizing, analyzing, and taking note of phenomena in transcribed speech' (MacWhinney, 2000:17), because it allows the researcher to include additional information that is based on his/her metalinguistic knowledge – an exclusively human ability.

Overall, such computer software enhance automation of analytic processes, such as computations on frequency of occurrence of a pattern; they increase reliability of data analysis by facilitating multiple assessments, both single researcher and inter-rater; and they ease subsequent sharing of data among researchers. In clinical assessment, in particular, computer technology may play a fundamental role in: a) the identification of errors and their nature, i.e. whether problems lie in phonemic identification, phonetic mapping or motor skill; b) optimizing intervention methods; and c) documenting and monitoring performance during treatment (Masterson & Rvachew, 1999; Bunta, Ingram & Ingram, 2003). The majority of these programs (and their manuals and usage tutorials) are freeware and available to the community online.

CLAN: A widely-recognized, though by no means exclusively used, language analysis software program is CLAN (MacWhinney, 2000) that offers separate tools for implementing transcription i.e. Codes for the Human Analysis of Transcripts (CHAT), and for analyzing language samples i.e. Computerized Language Analysis (CLAN), also allowing linguistic coding and systems for linking transcripts to audio and video excerpts. The program is designed to expedite a range of linguistic analyses, including phonological ones, and one of its strengths is that it enables several frequency computations, such as on words, morphemes, syllables, phonemes or codes. Data elicitation directly into CLAN is not facilitated, however. CLAN is the backbone of CHILDES, the Child Language Data Exchange System, (MacWhinney & Snow, 1985) that is currently the largest repository of children's spoken and written language, and a major linguistic resource. Detailed accounts and links to the program and the corpora (both monolingual and bilingual) can be found in the CHILDES website: <http://childes.psy.cmu.edu>. Although mostly associated with child language data in particular, CLAN can be used for linguistic analyses of speech in all its forms.

SALT: The Systematic Analysis for Language Transcripts program (Miller, Andriacchi & Nockerts, 2011; Heilmann & Malone, 2014) has been largely utilized by speech therapists. Like CLAN, it is designed for measuring all levels of language at the same time: morphology, syntax, semantic, pragmatic. The software permits the elicitation, transcription, coding, and analysis of language samples. It includes normative 'profiles' as a tool for the comparison and clinical management of children with language impairments. By comparing language samples (from conversation, narration, or exposition) to those of age-matched peers and generating Standard Measures reports, SALT identifies disordered performance, which may lead to language intervention strategies. It has built-in support for English, Spanish and French, but it can be used with many other languages. SALT, its manual and training sections can be found at www.saltsoftware.com.

CP: Computerized Profiling, (or Computerized language and phonological sample analysis (CL/PSA) (Long, Fey & Channel, 1998; Long & Channel, 2001) also enables complicated analyses that would not be feasible without the use of technology. The set of CP programs, available at <http://www.computerizedprofiling.org/downloads.html>, provides a wide range of linguistic measures (for grammar, semantics, pragmatics, narrative), but it differs from the aforementioned software in that it provides a comprehensive focus on phonological analysis. It enables computations of Percentage Consonant Correct (PCC), Phonological Mean Length of Utterance (PMLU), Proportion of Whole Word Proximity (PWP), Syllable Structural Level, and various independent and relational analyses (on segments, words, and prosody) (Ingram, 1981; Crystal, 1982; Shriberg & Kwiatkowski, 1982; Grunwell, 1987; Ingram & Ingram, 2001). Like CLAN, CP also performs various computations on structural statistics, such as utterance counts, etc.

CAPE: The Computerized Articulation and Phonology Evaluation System (Masterson & Bernhardt, 2001) is specifically designed for evaluating articulation and phonology from ages two (2) through adulthood and largely targets clinical assessment (<http://images.pearsonclinical.com/images/PDF/CAPE.pdf>). It enables time alignment of speech samples (either pre-entered as files or recorded on the spot) to orthographic text, automatically transcribes it into phonetics, and carries out phonological evaluation on the sample (e.g. computation of inventory and mismatch patterns), by directly comparing it to age-appropriate standards. Principally, CAPE carries out assessments in the English language, and includes some dialect filters (e.g. Spanish-influenced English).

PHON: The Phon program (Rose et al., 2006; Rose & MacWhinney, 2014), which is the successor of ChildPhon (Rose, 2003) enables analysis of speech data, both in transcript and acoustic form. It facilitates several tasks such as sound-to-text alignment, data segmentation,

data labeling and analysis, comparisons between targeted and produced forms, a powerful search system for data queries, multiple inter-rater transcription, links to multimedia, interoperability with Praat for acoustic analyses, a record navigator, etc. Like CAPES, Phon is designed exclusively for work on phonology and phonological development. But rather than targeting SSD assessment, Phon specializes as a tool for research in any aspect of phonology/phonetics in first language (L1), second language (L2), and bilingual (L1'2) speech, both normal and disordered. The software supports the PhonBank database, which provides shared speech records for the study of phonological development.

This list, though not exhaustive, presents some of the more established tools for phonetic and phonological analysis, though more recent efforts in this direction also exist, like EXMARaLDA (Schmidt & Wörner, 2009), ANVIL: the video annotation research tool (Kipp, 2014) and Web-Based Archiving (Tchobanov, 2014). The list excludes also technologies used for remediation (Kazakou et al., 2011).

Speech Analysis Software (SAS): working on acoustics

The accuracy and reliability of phonetic annotation as done by humans has been questioned in the literature due to its subjective nature (Gut & Bayerl, 2004; Hayward, 2014). Macken (1980:154), for instance, provides evidence of adult “listening bias” in that adult categorical perception might conceal sub-phonemic contrasts evident in children’s productions, such as the acoustically tested voicing contrast in the short-lag region. That perception might be a detriment to identifying speech sounds exactly is well-attested in the literature (Flege, 1991; Best, 1995; Blevins, 2007; Edwards & Beckman, 2008). It is generally recommended that phonetic transcription is supplemented by acoustic analysis of the transcribed data. There are a number of computer programs that are used for manipulating, visualizing, and analyzing the sound wave. They all have user-friendly interfaces and sound files can either be uploaded or recorded directly into the programs. Among these are:

Praat: A widely used software, Praat (Boersma & Weenink, 2015) permits both speech analysis and synthesis, labeling and segmentation, manipulation, computation of statistics on the data (e.g. multidimensional scaling, principal component, discriminant analyses), as well as the creation of graphics on selected speech signals. Speech and spectrum analysis provide information on several features of the signal like pitch, formant, intensity, jitter, etc. Praat may also be used for listening experiments to perform identification and discrimination tests. It is downloadable at www.praat.org.

SIL Speech Analyzer: This software provides an alternative to Praat, though with fewer features, facilitating fundamental frequency and spectrum analyses as well as duration measures. Orthographic, phonemic, tone and gloss annotations can supplement the phonetic transcription, while allowing easy access to the sound files.

Wavesurfer: Other than speech sound analysis and transcription, Wavesurfer may provide a platform for more advanced applications by extending it with new custom plug-ins or by embedding its components in other applications. It can serve as a tool for multiple tasks in speech research and in education, and targets beginners and advanced users.

In conclusion, there is a trend towards interoperability of existing programs that allows trouble-free transfer of data between them, such as that between CLAN-Phon (MacWhinney, 2012; Rose, 2014 pers. comm.), and across digital applications for additional analyses, i.e. acoustic (such as the interoperability of Phon and Praat), quantitative (the interoperability of CLAN with Microsoft Excel), or statistical like SPSS. A detailed presentation by Llisterri (2015) classifies the types of software related to speech analysis and transcription and includes links to the specific websites.

Intrusive technologies

Intrusive experiments require the physical involvement of participant/speakers as when, for instance, a part of the diagnostic equipment (e.g. a transducer) is in direct contact with them; these technologies, largely task-specific, focus on investigating motor, perceptual, and cognitive mechanisms in speech production that cannot be observed with the naked eye. The majority of data resulting from intrusive technologies involve pictorial representations (static 2-D, 3-D, 4-D, or video) of speech actions. Among technologies investigating speech motor mechanisms, such as the behavior of active and passive articulators (e.g. tongue, hard-palate), of the vocal cords, and of the larynx are:

- Electropalatography/EPG
- Ultrasound tongue imaging/UTI
- Magnetic resonance imaging/MRI
- Photoelectric glottography
- Electroglottography
- Inverse filtering
- Electromyography.

Perceptual and cognitive speech mechanisms giving insights into latent speech processing and brain activation are investigated using:

- MRI: Functional (fMRI) and diffusion (dMRI)
- Positron emission tomography
- Functional near-infrared spectroscopy/fNIRS.

Utilization of such technologies requires specialized training, and the experiment can only take place in laboratories carrying the relevant equipment. Comprehensive discussions on these technologies, which is beyond the scope here, may be found, among others, in Hardcastle & Hewlett (1999), Stemmer & Whitaker (2008), Park (2008), Malinen & Palo (2009), Simmonds, Wise & Leech (2011), Celata & Calamai (2012), Cohn, Foucheron & Huffman (2012), Ferrari & Quaresima (2012), Hardcastle, Laver & Gibbon (2012); Herff et al. (2012), Maassen & van Lieshout (2012), Scobbie, Stuart-Smith & Lawson (2012), and Hayward (2014).

NITs in use: a child case-study

This section provides an illustration of how NITs support investigation of large data-sets. Specifically, CLAN is utilized for the organization, management, retrieval, phonological and quantitative analysis of a female child's daily speech during her phonological development in simultaneous Greek/English bilingualism, longitudinally over a period of seventeen months (2;7-4;0); the main focus is on the development of consonantal segments. A database in CHAT format is created which includes alignment of the digitally recorded speech to orthographic and IPA transcriptions of the child's full utterances in both languages. Also, the processes observed in the child's consonant productions are coded in parallel with transcriptions of both targeted and produced speech. This is the longest spanning case-study of a child's phonological development with a focus on quantitative analysis of speech.

The CLAN database: data management and organization

The child's speech was digitally recorded on an Olympus WS-311M recorder (.wma audio files) on an almost daily basis from the age of 2;7 to 4;0. The database created includes all recorded audio files and respective CHAT transcripts. Following conversion into CLAN-compatible .wav format with Switch Sound File Converter, the converted audios of variable duration comprise the CLAN default Media folder, in layers of subfolders marking the difference in years and months. The database contains a total of 511 CHAT files with utterances in both languages, as they occurred in situ during naturalistic interaction with her mother. Each CHAT file (e.g. 2;7.05 WS310027) corresponds to an individual audio-file and is named to reflect the child's age on the day of recording (e.g. 2;7.05) and the standard name of the Olympus audio-file (e.g. WS310027). Thus, the structure of the CHAT (transcripts) folder mirrors that of the Media (audios) folder.

The default CHAT file is blank. Typed-in transcripts in CHAT files are preceded by headers (some are obligatory) that provide general information on the languages spoken, the participants, date(s) of the audio and CHAT file(s), the child's age, the transcriber's name, the interaction type, location, micro-contextual information, etc. All chat files in the database contain three identifiably main lines providing orthographic annotation: two for utterances corresponding to the child: *CHI (English), *CHIL (Greek), and one to the mother: *MOT (English). Though *CHI is the CLAN default tier, *CHIL was devised for the child's Greek utterances. Every *CHI or *CHIL tier is complemented by three subsequent tiers: the %mod, which provides a IPA transcription of targeted (model) speech; the %rep, which provides the IPA transcription of the child's produced (replica) speech; and %pho, the CLAN default tier for phonetic transcription which is used in this study for coding the child's productions. Thus, the %pho tier matches the model and replica tiers using a specific coding scheme, to be elaborated below. A typical child utterance (e.g. *CHI or *CHIL) with its phonetic transcripts and coding looks like a four-tier cluster (Table 1).

The coding scheme on the %pho tier follows this logic: the transcription on the %mod tier is repeated onto the %pho tier. Any consonant differences between the %mod and %rep tiers, including substitutions, deletions, epenthesis or syntagmatic processes (e.g. assimilations, coalescences) are noted following the %mod target segment. To exemplify this, the word *inside* from the utterance in Table 1 is used. The coded form '**in-sardT**' (%pho) in essence repeats /insard/ (%mod) and interprets the variants in the child's productions [ɪsart] (%rep) as follows:

n- i.e. /n/ → Ø, **dT** i.e. /d/ → [t]

Table 1. Example of a child utterance in English

*CHI: come inside the room
%mod: k ^h Λm insaid ðə ɹu:m
%rep: Λm ɪsaɪt ə lum
%pho: k ^h -Λm in-sardT ð-ə ɹLu:m

Table 2. Key to the coding scheme

IPA symbols and their Codes																								
IPA	p	b	f	v	t	d	s	z	θ	ð	ts	dz	l	r	m	n	j	c	w	k	g	x	ɣ	h
code	P	B	F	V	T	D	S	Z	Θ	Δ	TS	DZ	L	R	M	N	J	C	W	K	G	X	Γ	H
IPA	ɹ	ʃ	ʒ	tʃ	dʒ	ʌ	ŋ	ç																
code	R̥	S̥	Z̥	TS̥	DZ̥	L̥	N̥	C̥																
IPA	ɫ	ɲ	ɟ																					
code	Ł	N̄	J̄																					
number	assimilation					coalescence					metathesis													
	0					5					6													
other	deletion					reduplication					epenthesis													
	-					=					;													

The codes chosen in this study are symbols easily available on the keyboard. So, instead of Ø to mark deletion, a plain dash ‘-’ is used. In general, capitals stand for the respective segments, e.g. a [t] realization of any target (/.../→[t]) is coded as [T]. This coding applies to IPA symbols p, b, f, v, t, d, s, z, ts, dz, l, m, n, j, c, w, k, g, h that may be easily capitalized. By the same token, IPA symbols that find equivalence in the Greek alphabet, i.e. θ, χ, γ, are coded with their respective Greek capitals, e.g. Θ, Χ, Γ.

The voiced interdental ð is also coded using the Greek capital, Δ, that represents the sound. The flap, ɾ, in Greek is coded with capital R. The alveolar approximant, ɹ, in English is coded as R̥ using R and a randomly chosen IPA diacritic, ̥. Similarly, ʃ, ʒ, tʃ, dʒ, ʌ, ŋ and ç are coded as S̥, Z̥, TS̥, DZ̥, L̥, N̥ and C̥, respectively. The dark lateral, ɫ, is coded as Ł, (L and the IPA diacritic ~) because the final visual effect is similar. The same holds for the Greek palatals ɲ and ɟ coded as N̄ and J̄, respectively. Syntagmatic processes, like assimilation, coalescence, metathesis were also coded using numeric symbols: 0 (assimilation), 5 (coalescence), 6 (metathesis). Finally, symbols are used for coding deletion: ‘-’, reduplication ‘=’, and epenthesis ‘;’. An example of coded deletion was shown in Table 1. Vowels, overall ignored for the purposes of the study, are in broad transcription. A key to the codes used is given in Table 2.

The CLAN database: data retrieval and categorization

Once the data are entered in CLAN, the process of data retrieval becomes automated. Default CLAN commands target retrieval of specific information. Several commands were used to elicit general and specific information. For instance, the following command enables the tabulation of total utterances produced by the child in both languages: mlu+t*CHI-t%mor@. Also, a command was run per language on the total number of words used to determine vocabulary status in the two languages. The command for the word production in English was: freq+t*CHI-t*CHIL@. A snapshot of the outcome of this command is presented in Table 3. To track the developmental path of consonantal segments longitudinally, the following specific analyses were used to tabulate frequencies of:

Targeted segments (%mod): The following command, phonfreq+b%mod-t*CHIL+s"b"@, was used to tabulate how many times a segment, e.g. /b/, was targeted. The command is run separately for each targeted segment in the child’s languages. The command for targeted segments in the child’s Greek is slightly modified and run separately: phonfreq+b%mod-t*CHI+t*CHIL+s"b"@. A quick glimpse into the output files (usually extensive) for e.g. targeted /b/ is given in Table 4.

Produced segments (%rep): The following command tabulates e.g. the child’s [p] realizations (those correct in context, substitutions, and epenthetic ones): phonfreq+b%rep-

t*CHIL+s"p"@. Specifically, to tabulate the frequency of correct-in-context [p], `modrep+t*CHI-t*CHIL+b%mod+c%rep+o"*p"+n"*p"` is used instead. A section of this command's output can be seen in Table 5.

Substitutions produced: The command: `freq+t*CHI-t*CHIL+t%pho +s"*p"` is used to first identify substitutions corresponding to an individual segment, e.g. /p/ in English. This was done running the *freq* command on the %pho tier, identifying the particular segment examined and adding a specification for sensitivity to upper-case letters. The resulting output pairs the targeted segment with its substitution. The command is run separately in each language. A brief example section from these output files is found in Table 6.

Table 3. Command output for word frequency

```
freq +t*CHI -t*CHIL@
Mon Jul 30 13:10:16 2012
freq (28-Mar-2012) is conducting analyses on:
ONLY speaker main tiers matching: *CHI;
*****
From file <f:\ebab Database\CHAT files\2009\4.April 2009\2;7.05 WS310027_D.cha>
1 I
1 bowl
1 box
1 cereal
1 inside
1 milk
2 now
1 want
-----
8 Total number of different item types used
9 Total number of items (tokens)
0.889 Type/Token ratio >97
```

Table 4. Command output for targeted segments

```
phonfreq +b%mod -t*CHIL +sb @
Tue May 01 09:00:12 2012
phonfreq (28-Mar-2012) is conducting analyses on: ALL speaker main tiers EXCEPT the ones
matching: *CHIL; and those speakers' ONLY dependent tiers matching: %MOD;
*****
From file <e:\ebab Database\CHAT files\2009\4.April 2009\2;7.06 WS310031_D.cha>
*** File "f:\EBAB DATABASE\CHAT FILES\2009\4.April 2009\2;7.05 WS310010_A.cha": line 22
19 b initial = 17, final = 0, other = 2
```

Table 5. Command output for segment realizations

```
modrep +t*CHI -t*CHIL +b%MOD +c%REP +o"*p" +n"*p" @
Sun Feb 12 14:51:52 2012
modrep (02-Oct-2011) is conducting analyses on: ONLY speaker main tiers matching: *CHI; and
those speakers' ONLY dependent tiers matching: %MOD; %REP;
*****
From file <e:\EBAB DATABASE\CHAT FILES\2009\4.April 2009\2;7.06 WS310034_D.cha>
1 phIk
1 pit
From file <e:\EBAB DATABASE\CHAT FILES\2009\4.April 2009\2;7.09 WS310036_D.cha>
1 thop
1 thop
```

Table 6. Command output for segment substitutions

```

freq +t*CHI -t*CHIL +t%pho +s"*p*" @
Tue May 01 09:49:24 2012
freq (28-Mar-2012) is conducting analyses on: ONLY speaker main tiers matching:
*CHI; and those speakers' ONLY dependent tiers matching: %PHO;
*****
From file <e:\ebab Database\CHAT files\2009\4.April 2009\2;7.06 WS310031_D.cha>
1 ipB'ʊf
From file <e:\ebab Database\CHAT files\2009\4.April 2009\2;7.11 WS310041_D.cha>
1 pTh=ikT

```

Developmental and acoustic analysis results: the child's labio-velar approximant, w

As seen, CLAN outputs provide numerical computations on the data; these are further tabulated in Excel for conclusive reports. Results of tabulations, as shown above, reveal that in 135 hours of audio, there are 31,684 utterances containing 137,000 word tokens in the child's English data; data in Greek are comprised of 13,940 utterances containing 69,289 word tokens. In total, there are more than 200,000 child word-tokens in the entire database. The analysis above targets investigation of developmental paths of all consonants in the child's languages; due to space limitations, however, results are exemplified here by the development of the word-initial labio-velar approximant, w, in English. The corresponding targeted word-types in the child's English according to age of first appearance are:

2;7:	one /wʌn/, wait, walk, wall, want, water, way, we, wearing, wee, when, where, white, why, will, Willie, window, Winkie, Winnie, with, won't
2;8	ones /wʌnz/, warm, wash, watch, what, which, wide, windows, wolf
2;9	wake, wanted, went, wet, wheels, where, while, wind, work
2;10	waiting, wants, washed, watching, watermelon, weak, well, wipe, wonderful
2;11	water, waves, wheel, word, would
3;0	wagging, wagons, walks, wow, woke, working, worth
3;1	walking, wardrobe, washing, web, windy, winter, wiping, women, worker, workers
3;2	wagon, wave, week, weeks, whiskers, wing
3;3;	was, worm
3;4	warmed, Wednesday, were, whisker, whispering, win, wine, wish, wolves, wore, worry
3;6	weekend, welcome, whisper, without, wonder
3;7	waiting, Walrus, wetting, what's, whenever, wider, wind (to)
3;8	waking, warmer, wasn't, wiped
3;9	wished, wizard, wolfie
3;10	wedding, wicked, wild, wins
3;11	-

There are 112 word types (10,868 tokens). The monthly tokens are: 70, 139, 510, 513, 498, 567, 945, 777, 691, 965, 692, 854, 787, 1170, 796, 390, and 504. The development of the acquisition level of word-initial /w/ is shown in bold solid line in Figure 1. At 2;7 and 2;8, the level is lower than 20%. Between 2;9 and 3;5 the acquisition level fluctuates at about 55%. Starting at 3;6, it surpasses the 65% level while between 3;7 and 3;9 it is stationary just below the 90% level. At 3;10, /w/ is acquired completely at 97%. These results show a nine-month delay in this child's acquisition of /w/, since the monolingual English norm for acquisition of word-initial /w/ is by age 3;0 (Smit et al., 1990). A detailed discussion of the reasons for such a delay is beyond the purpose of this paper, however. The developmental substitutions of word initial /w/ whose frequency of occurrence in proportion to all substitutions at any age is higher than 10% are also shown in Figure 1. It is seen that the dominant substitution during development is [v] except near complete acquisition, when deletions and the vowel [ʊ] take over. The example evidently shows that retrieving specific information from audio

files and performing numerical computations that focus on a single segment at a time, e.g. the child's word-initial /w/-productions, would not have been reliable, if not physically impossible, without the use of technology, such as CLAN.

Additionally, the use of technology for acoustic analyses of speech segments confirms the reliability of phonetic transcription. Following analysis in Praat, example waveform and wideband spectrograms of *one* /wʌn/ and *wing* /wiŋ/ in the child's speech are shown in Figures 2 and 3.

Specifications in Praat were set at a window length of 5ms, a frequency range between 0-10 kHz for one and 0-5 kHz for wing, and the dynamic range at 30dB for one and 60db for wing. The acoustic characteristics of the child's [v] and [w] productions match those in respective targeted speech (e.g. Nirgianaki, 2014; Ladefoged & Johnson, 2015). The /w/→[v] substitution (Figure 2) is exemplified by overall low formants, typical of labial articulation, frication noise at over 8 kHz, and a light voicing bar; the striations in the pink section of the waveform also point to the fricative. The child's produced [w] (Figure 3) typically glides into the following vowel with F1, F2<1 kHz, and F3 slowly becoming clearer in transition to the following vowel at around 2 kHz.

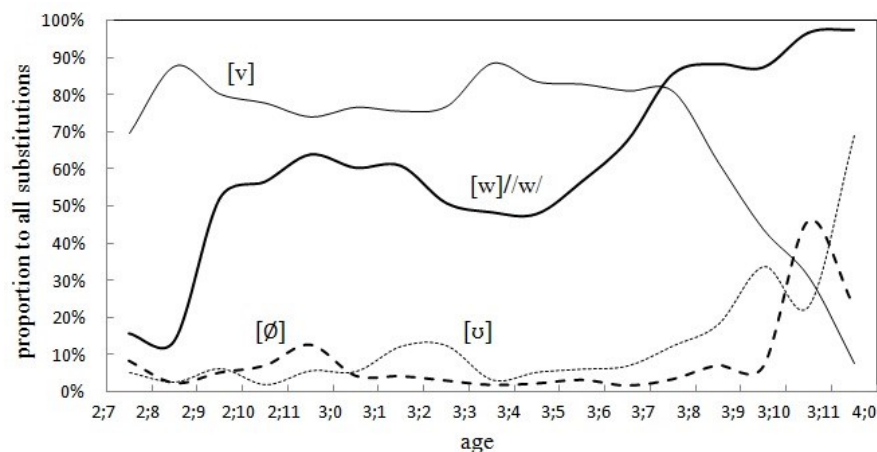


Figure 1. The development of word initial /w/

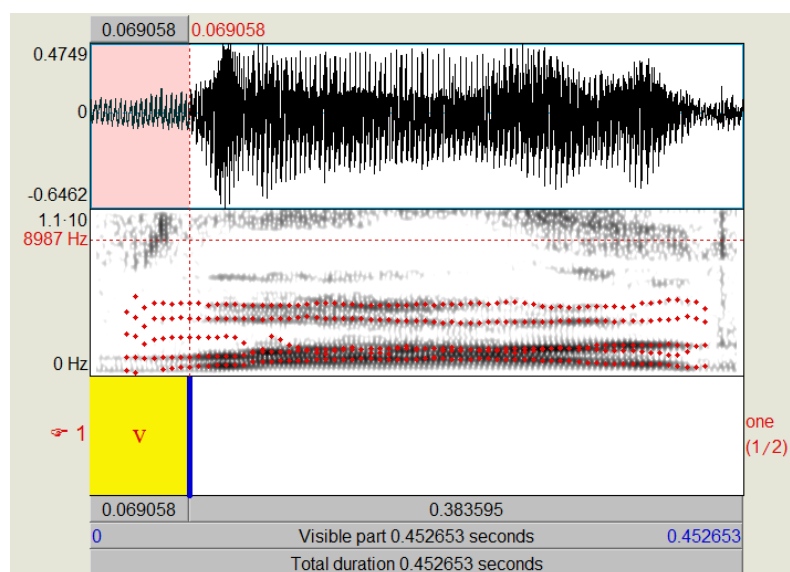


Figure 2. Example spectrogram of *one*, /w/→[v], in the child's speech

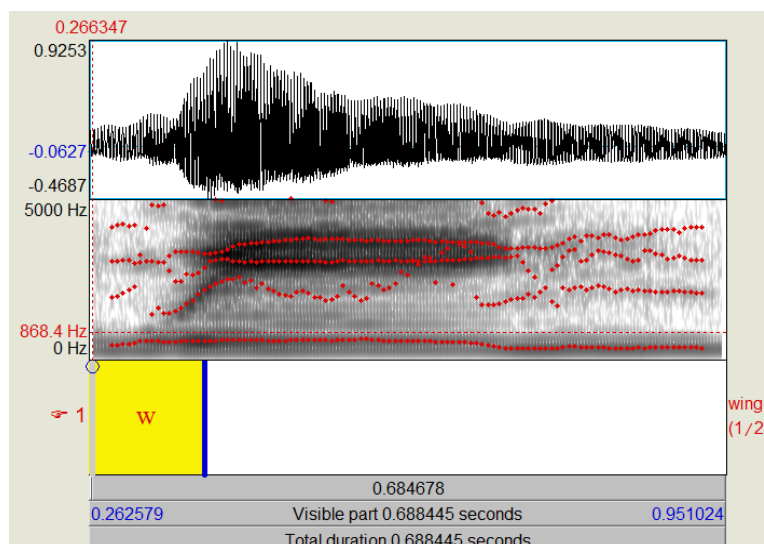


Figure 3. Example spectrogram of *wing*, /w/→[w], in the child's speech

Conclusions

This paper reviewed non-intrusive and intrusive technologies used for the study of speech. Though the review included a short discussion on intrusive technologies that investigate motor and cognitive aspects of speech by relying on the speaker's active participation in the experiment, the focus has been on non-intrusive technologies, referring to those technological advances that facilitate linguistic (phonological and phonetic) investigations of the speech signal in all its likely forms (adult, child, L1, L2, bilingual), normal or disordered. The technologies cited in the review are requisite for managing, retrieving, and performing various linguistic analyses and computations on small and large speech data-sets. This was illustrated by providing an example application: CLAN and Praat were utilized for investigating a child's speech data in a case study of Greek-English bilingualism, longitudinally over 17 months of phonological development. Specific results on the development of the labio-velar approximant, w, in the child's English are provided.

References

- Abercrombie, D. (1967). *Elements of general phonetics*. Edinburgh: Edinburgh University Press.
- Allen, W. (1953). *Phonetics in Ancient India*. London: Oxford University Press.
- Ball, M. J. (2008). Transcribing disordered speech: By target or by production? *Clinical Linguistics and Phonetics*, 22(10-11), 864-870.
- Ball, M. J. (2015). Multilingualism and acquired neurogenic speech disorders (plenary paper). In E. Babatsouli & D. Ingram (eds.), *Proceedings of the International Symposium on Monolingual and Bilingual Speech 2015* to appear.
- Ball, M. J., Müller, N., Klopfenstein, M., & Rutter, B. (2009). The importance of narrow phonetic transcription for highly unintelligible speech: Some examples. *Logoped Phoniatr Vocology*, 34(2), 84-90.
- Barras, C., Geoffrois, E., Wu, Z., & Liberman, M. (2001). Transcriber: Development and use of a tool for assisting speech corpora production. *Speech Communication*, 33(1), 5-22.
- Bell, A. M. (1867). *Visible speech: The science of universal alphabetic*. London: Simkin, Marshall & Co.
- Best, C. T. (1995). A direct realist view of cross-language speech perception. In W. Strange (ed.), *Speech perception and linguistic experience: Issues in cross-language research* (pp. 171-204). Timonium, MD: York Press.
- Blevins, J. (2007). Interpreting misperception: Beauty is in the ear of the beholder. In M. Solem, P. Bedor & M. Ohala (eds.), *Experimental approaches to phonology* (pp. 144-154). Oxford: Oxford University Press.
- Boersma, P., & Weenink, D. (2015). *Praat: Doing phonetics by computer* [Computer software]. Retrieved 25 April 2015, from <http://www.praat.org>.
- Bunta, F., Ingram, K., & Ingram, D. (2003). Bridging the digital divide: Aspects of computerized data collection and analysis for language professionals. *Clinical Linguistics and Phonetics*, 17, 217-240.

- Celata, C., & Calamai, S. (2012). Introduction. *Rivista di Linguistica / Italian Journal of Linguistics (special issue on articulatory techniques for sociophonetic research)*, 24(1), 1-9.
- Cohn, A. C., Fougerson, C., & Huffman, M. (eds.) (2012). *The Oxford handbook of laboratory phonology*. Oxford: Oxford University Press.
- Cole, J., Hasegawa-Johnson, M., Loehr, D., Van Guilder, L. Reetz, H., & Frisch, S. A. (2011). Corpora, databases, and Internet resources. In A. C. Cohn, C. Fougerson & M. K. Huffman (eds.), *The Oxford handbook of laboratory phonology* (pp. 441-458). Oxford: Oxford University Press.
- Crystal, D. (1982). *Profiling linguistic disability*. London: Edward Arnold.
- Cucchiaroni, C., Neri, A., & Strik, H. (2009). Oral proficiency training in Dutch L2: The contribution of ASR-based corrective feedback. *Speech Communication*, 51(10), 853-863.
- Duckworth, M., Allen, G., Hardcastle, W., & Ball, M. J. (1990). Extensions to the International Phonetic Alphabet for the transcription of atypical speech. *Clinical Linguistics and Phonetics*, 4, 273-280.
- Durand, J., Gut, U. & Kristoffersen, G. (2014) (eds.). *The Oxford handbook of corpus phonology*. Oxford: Oxford University Press.
- Edwards, J. G. H., & Beckman, M. E. (2008). Methodological questions in studying consonant acquisition. *Clinical Linguistics and Phonetics*, 22(12), 937-956.
- Edwards J. G. H., & Munson, B. (2012). Transcribing the speech of multilingual children with speech and sound disorders. In S. Macleod & B. Goldstein (eds.), *Multilingual aspects of speech sounds disorders in children* (pp. 170-181). Bristol: Multilingual Matters.
- ELAN (2015). *Linguistic annotator. Language archiving technology portal*. Nijmegen: Max Planck Institute for Psycholinguistics. Retrieved 25 March 2015, from <http://tla.mpi.nl/tools/tla-tools/elan>.
- Ferrari, M., & Quaresima, V. (2012). A brief review on the history of human functional near-infrared spectroscopy (fNIRS) development and fields of application. *Neuroimage*, 63(2), 921-935.
- Flege, J. E. (1991). Perception and production: The relevance of phonetic input to L2 phonological learning. In T. Huemmer & C. A. Ferguson (eds.), *Cross currents in second language acquisition and linguistic theories* (pp. 249-269). Amsterdam/Philadelphia: John Benjamins.
- Grunwell, P. (1987). *Clinical phonology*. (2nd edition). Baltimore, MD: Williams & Wilkins.
- Gut, U. (2012). The LeaP corpus. A multilingual corpus of spoken learner German and learner English. In T. Schmidt & K. Wörner (eds.), *Multilingual corpora and multilingual corpus analysis* (pp. 3-23). Amsterdam: Benjamins.
- Gut, U., & Bayerl, P. S. (2004). Measuring the reliability of manual annotation of speech corpora. In B. Bell & I. Marlien (eds.), *Proceedings of speech prosody* (pp. 565-568). Nara, Japan.
- Hardcastle, W. J., & Hewlett, N. (eds.) (1999). *Coarticulation: Theory, data and techniques*. Cambridge: Cambridge University Press.
- Hardcastle, W. J., Laver, J., & Gibbon, F. E. (eds.) (2012). *The handbook of phonetic sciences*. Malden, MA: Wiley-Blackwell.
- Hayward, K. (2014). *Experimental phonetics: An introduction*. London: Routledge.
- Heilmann, J., & Malone, T. (2014). The rules of the game: Properties of a database of expository language samples. *Language, Speech, and Hearing Services in Schools*, 45, 277-290.
- Herff, C., Heger, D., Putze, F., Guan, C., & Schultz, T. (2012). Cross-subject classification of speaking modes using fNIRS. In T. Huang, Z. Zeng, C. Li & C. S. Leung (eds.) *Neural Information Processing* (pp. 417-424). Springer: Berlin Heidelberg.
- IMDI-corpora (2015). *Domain of IMDI Described Corpora*. Retrieved 16 March 2015 from <https://corpus1.mpi.nl/ds/asv/?0>.
- Ingram, D. (1981). *Procedures for the phonological analysis of children's language*. Baltimore, MD: University Park Press.
- Ingram, D., & Ingram, K. (2001). A whole-word approach to phonological analysis and intervention. *Language, Speech, and Hearing Services in Schools*, 32, 271-283.
- IPA (1999). *Handbook of the International Phonetic Association*. Cambridge: Cambridge University Press.
- IPS (2014). *Institute of Phonetics and Speech Processing, Bavarian Archive for Speech Signals*. Retrieved 16 March 2015 from <https://clarin.phonetik.uni-muenchen.de/BASWebServices/#/services>.
- Jakobson, R. (1941/1968). *Child language, phonological universals and aphasia*. (Keiler, A. Trans.). The Hague: Mouton. Original work published in 1941 as *Kindersprache, aphasie und allgemeine Lautgesetze*.
- Kazakou, M., Soulis, S., Morfidi, E., & Mikropoulos, T. A. (2011). Phonological awareness software for dyslexic children. *Themes in Science and Technology Education*, 4(1), 33-51.
- Kipp, M. (2014). ANVIL: The video annotation research tool. In J. Durand, U. Gut, & G. Kristoffersen (eds.) *Handbook of Corpus Phonology* (pp. 420-436). Oxford : Oxford University Press.
- Ladefoged, P., & Johnson, K. (2015). *A course in phonetics*. Stamford, CT: Cengage Learning.
- Llisterri J. (2015). Speech analysis and transcription tools. Retrieved 16 March 2015, from http://liceu.uab.cat/~joaquim/phonetics/fon_anal_acus/herram_anal_acus.html.
- Long, S. H., & Channel, R. W. (2001). Accuracy of four language analysis procedures performed automatically. *American Journal of Speech Language Pathology*, 10, 180-188.

- Long, S. H., Fey, M. E., & Channel, R. W. (1998). *Computerized profiling (CP)*. [Computer software]. Cleveland, OH: Department of Communication Sciences, Case Western University.
- Maassen, B., & van Lieshout, P. (2012). *Speech motor control: New developments in basic and applied research*. Oxford: Oxford Scholarship Online.
- Macken, M. A. (1980). Aspects of the acquisition of stop systems: A cross-linguistic perspective. In G. H. Yeni-Komshian, J. F. Kavanagh & C. A. Ferguson (eds.), *Child phonology 1: Production* (pp. 143-168). New York: Academic Press.
- MacWhinney, B. (2000). *The CHILDES project: Tools for analyzing talk*. Mahwah, NJ: Lawrence Erlbaum.
- MacWhinney, B. (2007). The TalkBank Project. In J. C. Beal, K. P. Corrigan & H. L. Moisl (eds.), *Creating and digitizing language corpora: Synchronic databases* (Vol. 1., pp. 163-180). Houndmills: Palgrave-Macmillan.
- MacWhinney, B. (2012). *Personal communication*. 20 October, SLRF 2012 Pittsburgh, PA.
- MacWhinney, B., & Snow, K. (1985). The child language data exchange system. *Journal of Child Language*, 12, 271-296.
- Malinen, J., & Palo, P. (2009). Recording speech during MRI: Part II. *MAVEBA 2009* (pp. 211-214). Florence, Italy.
- Masterson, J. J., & Rvachew, S. (1999). Use of technology in phonological intervention. *Seminars in Speech and Language*, 20(3), 233-250.
- Masterson, J. & Bernhardt, B. (2001). Computerized articulation and phonology evaluation system (CAPES). San Antonio, TX: Psychological Corporation.
- Miller, J. F., Andriacchi, K., & Nockerts, A. (2011). *Assessing language production using SALT software: A clinician's guide to language sample analysis*. Middleton, WI: SALT Software LLC.
- Nirgianaki, E. (2014). Acoustic characteristics of Greek fricatives. *Journal of the Acoustical Society of America*, 135(5), 2964-2976.
- Park, H. (2008). Brain activation patterns in perception of Korean laryngeal contrasts. *Korean Linguistics*, 14(257), 1-19.
- Rose, Y. (2003). ChildPhon: A database solution for the study of child phonology. In B. Beachley, A. Brown & F. Conlin (eds.), *Proceedings of the 27th Annual Boston University Conference on Language Development* (pp. 674-685). Cascadilla Press.
- Rose, Y. (2014). Personal communication (email to author).
- Rose, Y., MacWhinney, B., Byrne, R., Hedlund, G., Maddocks, K., O'Brien, P., & Wareham, T. (2006). Introducing Phon: A software solution for the study of phonological acquisition. In D. Bamman, T. Magnitskaia & C. Zaller (eds.), *Proceedings of the 30th Annual Boston University Conference on Language Development* (pp. 489-500). Somerville, MA: Cascadilla Press.
- Rose, Y., & MacWhinney, B. (2014). The PhonBank Project: Data and software-assisted methods for the study of phonology and phonological development. In J. Durand, U. Gut & G. Kristoffersen (eds.), *The Oxford handbook of corpus phonology* (pp. 308-401). Oxford: Oxford University Press.
- Ruhi, S., Haugh, M., Schmidt, S., & Wörner, K. (eds.) (2014). *Best practices for spoken corpora in linguistic research*. Cambridge: Cambridge University Press.
- Schiel, F., Draxler, C., & Harrington, J. (2011). *Phonemic segmentation and labelling using the MAUS technique*. Retrieved 16 March 2015 from <https://www.phonetik.uni-muenchen.de/forschung/publikationen/Schiel-VLSP2011.pdf>.
- Schmid, M. S., Verspoor, M., & MacWhinney, B. (2011). Coding and extracting. In M. Verspoor, K. de Bot & W. Lowie (eds.), *A Dynamic approach to second language development* (pp. 39-54). Amsterdam/Philadelphia: John Benjamins.
- Schmidt, T., & Wörner, K. (2009). EXMARaLDA - Creating, analysing and sharing spoken language corpora for pragmatic research. *Pragmatics*, 19(4), 565-582.
- Scobbie, J. M., Stuart-Smith, J., & Lawson, E. (2012). Back to front: A socially-stratified ultrasound tongue imaging study of Scottish English /u/. *Italian Journal of Linguistics*, 24(1), 103-148.
- Shriberg, L. D., & Kent, R. D. (2003). *Clinical phonetics*. Boston, MA: Allyn & Bacon (3rd edition).
- Shriberg, L. D., & Kent, R. D. (2013). *Clinical phonetics*. Boston, MA: Pearson Education (4th edition).
- Shriberg, L. D., & Kwiatkowski, J. (1982). Phonologic disorders III: A procedure for assessing severity of involvement. *Journal of Speech and Hearing Disorders*, 47, 256-270.
- Simmonds, A. J., Wise, R. J. S., & Leech, R. (2011). Two tongues, one brain: imaging bilingual speech production. *Frontiers in Psychology*, 2, 166.
- Sloetjes, H., & Wittenburg, P. (2008). Annotation by category-ELAN and ISO DCR. In N. Calzolari (ed.), *Proceedings of the 6th International Conference on Language Resources and Evaluation* (pp. 816-820).
- Smit, A. B., Hand, L., Freilinger, J. J., Bernthal, J. E., & Bird, A. (1990). The Iowa articulation norms project and its Nebraska replication. *Journal of Speech and Hearing Disorders*, 55, 779-798.
- Stemmer, B., & Whitaker, H. A. (2008). *Handbook of the neuroscience of language*. London, UK: Academic Press
- Strik, H., Truong, K., de Wet, F., & Cucchiaroni, C. (2009). Comparing different approaches for automatic pronunciation error detection. *Speech Communication*, 51(10), 845-852.

- Tchobanov, A. (2014). Web-based archiving and sharing of phonological corpora. In J. Durand, U. Gut, & G. Kristoffersen (eds.) *The Oxford Handbook of Corpus Phonology* (DOI: 10.1093/oxfordhb/9780199571932.013.029). Oxford: Oxford handbooks online.
- Wald, M., & Bain, K. (2008). Universal access to communication and learning: The role of automatic speech recognition. *Universal Access in the Information Society*, 6(4), 435-447.
- Weinberger, S. H. (2015). *Speech accent archive*. George Mason University.
- Weinberger, S. H., & Kunath, S. A. (2011). A computational model for accent identification. In M. Wrembel, M. Kul & K. Dziubalska-Kolaczyk (eds.), *Achievements and perspectives in the acquisition of second language speech: New Sounds 2010* (pp. 287-296). Frankfurt: Peter Lang.

To cite this article: Babatsouli, E. (2015). Technologies for the study of speech: Review and an application. *Themes in Science and Technology Education*, 8(1), 17-32.

URL: <http://earthlab.uoi.gr/thete/index.php/theste>